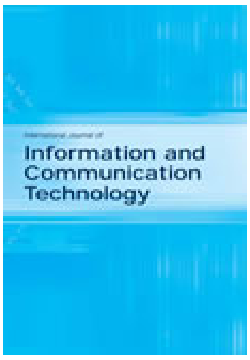
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Cross-border trade export prediction based on reinforcement learning and multimodal data

Yang Yu, Ziwen Sun, Jiachen Li

DOI: 10.1504/IJICT.2025.10073651

Article History:

Received:	24 May 2025
Last revised:	22 August 2025
Accepted:	22 August 2025
Published online:	10 October 2025

Copyright © 2025 Inderscience Enterprises Ltd.

Cross-border trade export prediction based on reinforcement learning and multimodal data

Yang Yu

School of Modern Finance,
Jiaxing Nanhu University,
Jiaxing, 314001, China
Email: yy111700@126.com

Ziwen Sun

National University Science Park,
Hangzhou Normal University,
Hangzhou, 311121, China
Email: ziwen511@sina.com

Jiachen Li*

School of Economics,
Fudan University,
Shanghai, 200433, China
Email: jupiter_jiachen@163.com
*Corresponding author

Abstract: Cross-border trade export forecasting is important for enterprises to optimise resource allocation. However, existing prediction methods have the problem of insufficient single modal feature extraction, for this reason, this paper first optimises the reinforcement learning (RL) algorithm based on multilevel strategy and multilevel reward (MSRL). Then CNN, Doc2Vec model, and improved ResNet152 model were used to extract static variable features, comment text features, and image features of cross-border trade export sales volume, respectively, and a hierarchical attention mechanism was designed to fuse multimodal features. The hyperparameters of the BiGRU model are optimised using MSRL (MSRL-BiGRU), and the fusion features are input into MSRL-BiGRU, which efficiently and automatically searches for the optimal strategy and reduces the prediction error. The experimental results show that the proposed method improves the coefficient of determination R^2 by 4.84–18.67%, which can realise the accurate prediction of cross-border trade export sales.

Keywords: cross-border trade export forecasting; reinforcement learning; multimodal fusion; hierarchical attention mechanism; BiGRU model.

Reference to this paper should be made as follows: Yu, Y., Sun, Z. and Li, J. (2025) 'Cross-border trade export prediction based on reinforcement learning and multimodal data', *Int. J. Information and Communication Technology*, Vol. 26, No. 36, pp.1–16.

Copyright © The Author(s) 2025. Published by Inderscience Publishers Ltd. This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Biographical notes: Yang Yu received his PhD in Fudan University in July 2017. He is now working at Jiaxing Nanhu University. His research interests include machine learning, digital economy, and green financial trade export.

Ziwen Sun received his PhD in Shikoku University of Japan in 2021. He is currently a minister of National University Science Park in Hangzhou Normal University. His research interests include digital economy and industrial intelligence.

Jiachen Li s currently a PhD candidate in the School of Economics at Fudan University. His research interests include digital trade and financial technology.

1 Introduction

Cross-border e-commerce refers to a commercial activity in which transaction subjects between different countries use e-commerce platforms to facilitate cooperative transactions, then pay for the amount of the purchase, and finally transport the goods through cross-border logistics, which is an inseparable part of the current cross-border trade field and occupies a considerable proportion (Wang et al., 2020). Cross-border e-commerce sales data directly reflect the characteristics of the flow of goods, and thus these sales data can be collected and used to make predictions about the future exports of cross-border trade (Wang et al., 2017). Because of the lag in the supply of goods, firms need to plan their trade exports as accurately as possible in the light of market demand over a period of time in the future, but future market demand is uncertain. If firms overestimate international market demand, they will incur inventory build-ups, which in turn will incur inventory expenses and capital costs (Saydam and Civelek, 2022). Therefore, accurate and efficient cross-border trade export forecasting is key for firms to reduce uncertainty and minimise inventory build-up and opportunity costs (Elia et al., 2021).

The cross-border trade export forecasting problem belongs to the field of time series forecasting. Classical statistical methods such as linear regression (LR) (Lourenço et al., 2011) and moving average autoregressive model (Pham and Yang, 2010) are widely used in this field. Kalaoglu et al. (2015) introduced traditional festivals as dummy variables into LR to predict future sales of foreign trade clothing based on the long-term and seasonal characteristics of the time series. Jiang et al. (2021) selected macro indicators such as total retail sales of consumer goods, GDP and primary industry output as variables and input them into a multiple regression model to predict logistics demand relatively accurately. Menculini et al. (2021) used a Prophet-ARIMA model to forecast wholesale food prices, resulting in poor predictive accuracy of the model.

Traditional statistical methods perform poorly in the prediction problem due to their poor predictive fitting ability and high requirement for completeness of historical data. However, as the artificial intelligence growing, deep learning models have a strong learning ability and are significantly more effective in prediction problems. Ouyang et al. (2019) used LSTM to extract impact indicator features to better enable the prediction of online agricultural prices, and the model was significantly better than a shallow classifier without feature learning. Massaro et al. (2021) proposed an e-commerce sales prediction model based on XGBoost, and compared the prediction accuracy and real-time performance with LSTM model to verify the effectiveness of the model. E-commerce data often presents a long-tail distribution, with very little data on some products or user behaviours, making it difficult for models to capture patterns and resulting in poor prediction effects. Pan and Zhou (2020) utilised CNN to extract features from review texts of e-commerce platforms and output the prediction results through a fully connected network, but the prediction accuracy was not high.

Deep learning-based forecasting methods focus on immediate outputs, making it difficult to directly optimise long-term goals. Reinforcement learning-based prediction methods learn the optimal strategy through the continuous interaction between the agent and the environment. In e-commerce sales prediction, the environment can be regarded as market dynamics, and the agent adjusts the parameters of the prediction model through trial and error. Through this adaptive interaction strategy, the prediction performance can be greatly enhanced. Lee et al. (2022) introduced a reward and punishment mechanism to optimise the prediction indicators, and constructed a ‘strategy network’ and a ‘value network’ to optimise the prediction results in the decoding stage to improve the prediction performance. An et al. (2023) used a BERT-CNN model to extract textual features of user comments on e-commerce platforms and predicted cross-border trade exports with a prediction accuracy of 78.2% through an RL-optimised LSTM model. The above prediction methods based on a single modality only consider source data features or text features, while cross-border e-commerce platforms have multimodal data, and how to better integrate these multimodal data has become a research hotspot. Li (2024) integrated image features in a multilayer perceptron (MLP) while utilising an RNN to capture the context and location relationships of commodity attributes to improve the performance of multimodal prediction models. Cai et al. (2021) used a cross-modal recurrent neural network based on the bimodal attention mechanism (CA-RNN) to predict merchandise sales from multimodal data. Xu et al. (2024) significantly improved the prediction accuracy by using CNN and BiLSTM for text and image data feature extraction, respectively, and multimodal feature fusion through the attention mechanism.

According to the comprehensive analysis of the above research status, it can be seen that the current cross-border trade export prediction method of single modal feature extraction is insufficient, resulting in large prediction errors, for this reason, this paper proposes a cross-border trade export prediction method based on the fusion of RL and multimodal data. The main work of the method is summarised in the following aspects:

1 Optimisation of RL algorithm based on the idea of multilevel strategy and multilevel reward (MSRL). The multilevel strategy module provides maximised confidence by jointly updating semantic strategies, and the multilevel reward network jointly evaluates visual-linguistic relevance through visual rewards to obtain RL-specific goals and improve global decision optimisation.

2 Static variable features, comment text features and image features of cross-border trade export sales were extracted using CNN, Doc2Vec model, and improved ResNet152 model, respectively. Designing hierarchical attention mechanisms to meticulously capture and fuse features from different modal data improves the alignment between different modal features and learns better quality feature representations.

3 The MSRL is used to optimise three hyperparameters, namely the amount of neurons, the amount of epochs, and the BiGRU model's studying rate, and the three parameters to be optimised are established as three states, each with a range of actions, and intelligent agents incessantly engage with the BiGRU in search of the best policy, so that the optimal cross-border trade export sales prediction results are obtained upon reaching the maximum allowed iterations.

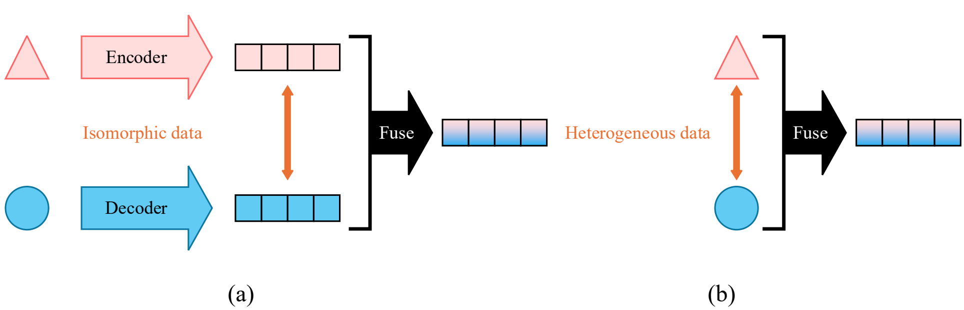
4 The experimental results show that the MAE and MAPE of the proposed method are reduced by 12.68–50.1% compared with the other three models, which greatly reduces the prediction error and improves the prediction performance, and has high practical value.

2 Relevant theoretical foundations

2.1 Theory of multimodal data fusion

The core problem of multimodal data fusion is to combine the information of various different modal data, which fuses multimodal data into a single feature in order to achieve comprehensive judgement and decision making on the same phenomenon (Gao et al., 2020). The purpose of multimodal representation fusion is to study joint representations that model the interactions of individual elements between various modalities, integrating information from two or more modalities and effectively reducing the amount of separate representations. Multi-modal fusion can be categorised into two main types: fusion of abstract modes (FAM) and fusion of primitive modes (FRM) (Zhang et al., 2020), as shown in Figure 1.

Figure 1 Multimodal fusion method, (a) fusion with abstract modalities (b) fusion with raw modalities (see online version for colours)



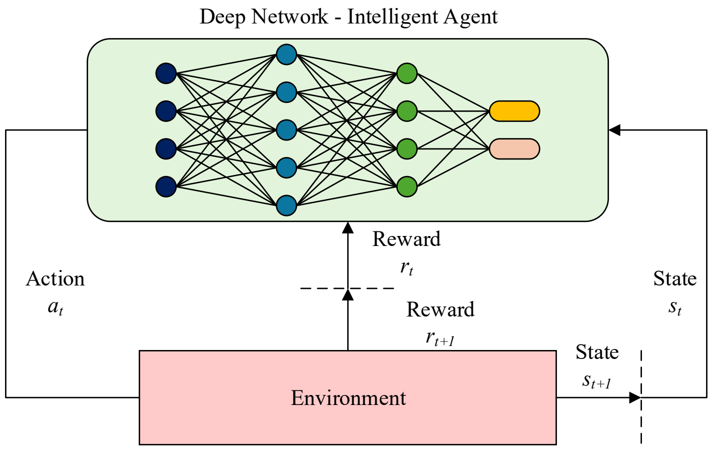
In FAM, an appropriate unimodal encoder is first used to obtain an overall representation of each element, and then the information from the two data streams is flexibly fused together using addable and multiplicable interactive operations that act as differentiable building blocks, as shown in Figure 1(a). These operations can be represented into almost any unimodal machine learning process.

In contrast, FRM occurs at a very early stage and may even involve the raw modal data itself, as shown in Figure 1(b). These methods are often similar to early fusion, where the input data are manipulated in series before the predictive model is applied.

2.2 Reinforcement learning algorithm

RL refers to learning by trial-and-error of continuous interaction between an intelligent (A) and its environment (E), which enables A to perform actions that receive maximum reward (Matsuo et al., 2022). In RL, E scores the actions performed by the intelligences, A learns the actions with the highest scores in each state, and finally A composes an optimal policy from the best actions in each state. At each time step t, A observes the current state s_t of the environment, makes an action a_t according to the established strategy, E feeds a reward message r_t to A based on the action made by A, and A chooses the next action based on r_t and s_t while E moves from s_t to the next state s_{t+1}.

Figure 2 The structure of deep RL (see online version for colours)



In reinforcement learning, denote all rewards from the initial to the end as $r_1, \ldots, r_t, \ldots, r_n$. Define the discount rate $\gamma \in [0, 1]$, then the discounted reward is expressed as follows.

$$U_t = r_t + \gamma \cdot r_{t+1} + \gamma^2 \cdot r_{t+2} + \ldots + \gamma^{n-t} \cdot r_n \tag{1}$$

At time t, U_t is an unknown random variable whose randomness comes from all the states and actions after time t. The action value function is expressed as follows. The action value function is expressed as follows:

$$Q_\pi(s_t, a_t) = E\left[U_t \mid S_t = s_t, A_t = a_t\right] \tag{2}$$

The expectation in the above formula eliminates all states after time t for all actions. The optimal action value function is expressed by maximising elimination strategy π as below:

$$Q_*(s_t, a_t) = \max_\pi Q_\pi(s_t, a_t) \tag{3}$$

RL has significant advantages in dynamic decision making, long-term optimisation, autonomous exploration, etc., and is suitable for scenarios that require interaction between agents and the environment. Deep learning is better at pattern recognition and processing static data. Combining the two into deep RL can leverage their respective strengths to solve more complex problems, as shown in Figure 2.

3 Optimisation of reinforcement learning algorithms with joint multilevel policy and reward networks

Traditional RL algorithms that consider only a single policy may prematurely converge to a locally optimal solution instead of a globally optimal solution. To this end, the joint multilevel policy and multilevel reward ideas are optimised for reinforcement learning algorithms (MSRL), where the multilevel policy module provides maximised confidence by jointly updating the semantic policies, and the multilevel reward network jointly evaluates the visual linguistic relevance by visual rewards to obtain RL-specific goals and improve the global optimisation searching ability. Visual-level strategies, as the perception core in multi-level strategies, form a collaborative optimisation relationship with multi-level strategies through hierarchical abstraction, shared state representation and joint training. Its design details, such as feature extraction, action space, and reward design, not only follow the general logic of multi-level strategies but also adapt to the particularity of visual tasks. This balance between similarity and difference enables the multi-level strategy framework to handle complex tasks efficiently while maintaining focus and interpretability at each level.

The multilevel policy network consists of semantic-level policies and visual-level policies. The semantic level policy embeds the features I by extracting them from the text using a linear mapping. The visual-level strategy $c(I, S_t)$ belongs to the visual embedding network. The multilevel reward network consists of a visual reward θ_r and a semantic reward S_t. θ_r is from the visual features while S_t is the semantic features. $\varpi = \{\theta_\pi, \theta_a\}$ denotes the parameters of the multilevel strategy network, θ_π and θ_a are the strategy network and reward network respectively and r_{total} is rewarded by minimising the negative expectation combinations. The sample approximation of MSRL is obtained by training ϖ together with θ_π and θ_r as shown below:

$$\nabla_{\theta_\pi} \xi \approx \sum_{t=1}^{T} \nabla_{\theta_\pi} \log p_\pi \left(w_t \middle| I, S_t \right) \tag{4}$$

where w_t is the word distribution. Considering this gradient scaling as an estimate a_t of the action advantage at state s_t, the value function of MSRL is defined as its expected future return, as shown below:

$$V_\theta \left(g_t \middle| x^n, y^n \right) = E_{g_t} \left[R \left(g_t \middle| x^n, y^n \right) \right] \tag{5}$$

where g_t is the reward function at a given moment, x^n and y^n are different modal data, the expectation is $g_t \sim \pi_\theta(\cdot|g_t, x^n)$, and π_θ is a multilevel strategy. The goal is to maximise the average reward from the initial state s_0 defined as follows:

$$J(\theta) = \frac{1}{N} \sum_{n=1}^{N} V_\theta \left(s_0 \middle| x^n, y^n \right) \tag{6}$$

where N is the number of samples in the training set, and only one example n is considered, omitting the numbers x^n and y^n. The gradient of $J(\theta)$ is computed. In the special case of deterministic transition functions, the gradient of $J(\theta)$ is as follows, where the Q-function is $Q_\theta(g_t) = E_{g_t}[R(g_t)]$.

$$\nabla_{\theta} V_{\theta}(s_0) = E_{g_t}\left[\sum_{t=1}^{T} \nabla_{\theta}\pi_{\theta}(g_t) Q_{\theta}(g_t)\right] \tag{7}$$

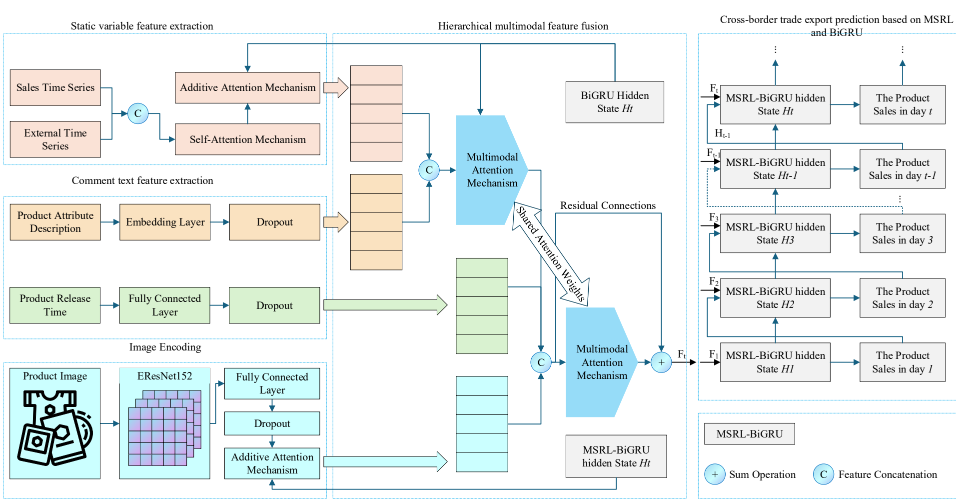
The goal of MSRL is to estimate the Q-function using Monte Carlo by first sampling all K samples of s_t and g_t, then computing the mean. Based on all the complete sequences sampled by the current multilevel strategy network, the optimal Q-function estimation is performed based on the multilevel reward function conditional on some of the sequences sampled in the current strategy to obtain the global optimisation result.

4 Cross-border trade export prediction based on reinforcement learning and multimodal data fusion

4.1 Multimodal cross-border trade export impact data feature extraction

To address the issue that the current cross-border trade export prediction methods have insufficient single-modal feature extraction, which leads to low prediction accuracy, the static variable features, comment text features and image features of cross-border trade export sales are first extracted respectively. Then a hierarchical attention mechanism is designed to fuse the features of the above modalities, and the hyperparameters of the BiGRU model are optimised by using MSRL (MSRL-BiGRU), and the fused features are inputted into MSRL-BiGRU, which efficiently and automatically searches for the optimal strategy, improves the efficiency of hyperparameter searching, and thus reduces the prediction error of BiGRU. The entire flow of the offered prediction method is shown in Figure 3.

Figure 3 The entire flow of the offered prediction method (see online version for colours)



Cross-border trade export sales in addition to domestic and foreign policies, product quality, market demand, marketing strategy and other intrinsic variables, but also by the online sales platform user's multimodal comment information, these comments are mostly text and image data, for this paper will use different algorithms for static variable features, comment text features and image features are extracted separately, the specific steps are as follows:

- Static variable feature extraction: firstly, the static variable field is converted into a numeric list and mapped into vector m_i using the embedding matrix, and finally m_i is inputted into the CNN, and the feature vector m_i' is obtained by utilising the fully connected layer with the ReLU activation function, as shown below, where w_1 is the weight and b_i is the bias.

  $$m_i' = ReLU\left(w_i^T m_i + b_i\right) \tag{8}$$

  After obtaining the features of all static variables, stitching is performed to finally obtain the static feature vector m_i^* of cross-border trade export sales.

- Product review text feature extraction: in order to accurately characterise the text, this paper adopts the Doc2Vec model (Kim et al., 2019), which generates a unique paragraph ID for each text, and maps sentences and paragraphs directly into a vector space of fixed dimensions. The Doc2Vec model consists of a distributed memory model (PV-DM) and a distributed bag-of-words model (PV-DBOW). Since PV-DM performs better than PV-DBOW, this paper adopts PV-DM to extract text features of product reviews.

  First, each text paragraph ID and context words are initialised into paragraph vector matrix $D = [d_1, d_2, \ldots, d_M]$ and word vector matrix $W = [o_1, o_2, \ldots, o_N]$, where each column represents a paragraph or a word, where $d, o \in R^{q\times 1}$, q is the dimension, M is the number of paragraphs, and N is the number of words. A window of radius x is slid over the sentences of a paragraph, and every time it is slid, d_i and the context word $(o_{j-x}, \ldots, o_{j+x})$ in the window are spliced together to obtain an aggregation vector, and finally the probability of the target word is obtained by softmax. The goal of model training is to maximise the average log-likelihood function, where the conditional probabilities are calculated as follows:

  $$p\left(o_j \middle| d_i, o_{j-x}, \ldots, o_{j+x}\right) = \frac{e^{z_{o_j}}}{\sum_g e^{z_g}} \tag{9}$$

  $$z = Uh\left(d_i, o_{j-x}, \ldots, o_{j+x}; W, D\right) + b \tag{10}$$

  where o_j is the current word vector; z_g is the non-normalised logarithmic probability of the g^{th} word, W is the sentence vector matrix, D is the word vector matrix, U and b are softmax parameters. The training is performed by W and U, b, and the passage vector of the new passage, i.e., the text feature vector f_i^*, is obtained by using the gradient descent method.

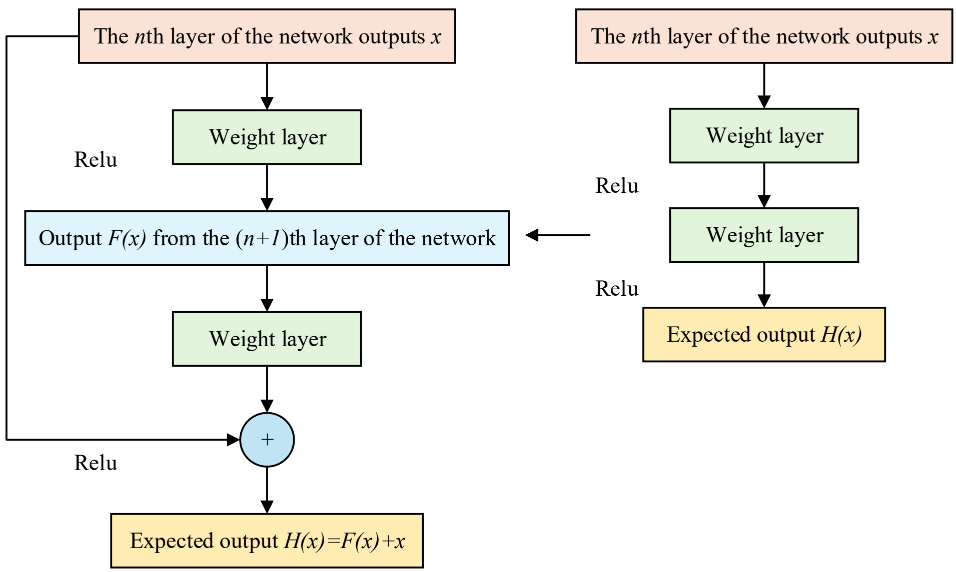
- Image feature extraction based on improved ResNet152, VGG-16, etc. are the more popular image feature extraction methods (Zhang et al., 2022), among which ResNet152 is widely used due to its excellent performance. For this reason in this paper, ResNet152 is improved (EResNet152) using the residual block mapping method, as shown in Figure 4.

The output x of layer $F(x)$ is added to the input $H(x) = F(x) + x$ of layer n to obtain the desired output D using the ReLU activation function. In the process of reverse gradient propagation, ensure that each expected value is added to the output value, thereby reducing the phenomenon of gradient dispersion. The feature extraction formula of EResNet152 is as follows:

$$h_i = f\left(W_i^T(\text{EResNet}(x)) + b\right) \tag{11}$$

where h_i is the image feature vector, vector size of 512 dimensions, W is a matrix of 2,048 × 512 dimensions, x is the image data input.

Figure 4 The residual block mapping method (see online version for colours)



4.2 *Hierarchical attention-based feature fusion for multimodal data*

Adopting a hierarchical feature fusion approach can help the model to obtain a richer semantic representation. Considering that the static variable features and textual features contain more abstract information, priority is given to the joint feature encoding of the m_i^* and f_i^* pairs. They are subsequently fused with low-level semantic modal image features to complement each other. In the field of cross-border trade and export, both high-level and low-level semantic information have a significant impact on the model's prediction results.

The multimodal attention (MMA) designed in this paper takes the joint feature representations $F_{a\&b}$ and h_i of m_i^* and f_i^* as inputs, which facilitates the model to focus directly on information with strong correlations between different modalities, thus improving the alignment between cross-modal data. In the multimodal fusion process, two MMAs share attentional weights to help model modal interdependencies, which are calculated for each MMA as shown below:

$$\begin{aligned} f_{MMA}\left(F_q, F_k, F_v\right) &= f_{softmax}\left(\frac{F_q F_k^T}{\sqrt{d}}\right) F_v \\ &= f_{softmax}\left(\frac{F_h W_q (F_{a\&b} W_k)^T}{\sqrt{d}}\right) F_{a\&b} W_v \end{aligned} \tag{12}$$

$$F_{a\&b} = f_{concat}\left(m_i^*, f_i^*\right) \tag{13}$$

where f_{MMA}, $f_{softmax}$ and f_{concat} are the MMA, softmax and concatenate functions, respectively; F_q, F_k, F_v is the query, key and value of MMA, respectively; $F_{a\&b}$ is the joint feature representation of m_i^*, f_i^*-connections and F_h is the feature representation of the image; S and D are the sequence length and feature dimensions, respectively, W_q, W_k and W_v are the learnable parameter matrices of the query, key and value, respectively.

The attention weights of F_q and F_k map the two vectors to scalars through the attention scoring function α to obtain the final multimodal feature fusion result F^*.

$$\alpha\left(F_q, F_k\right) = f_{softmax}\left(\frac{F_q F_k^T}{\sqrt{d}}\right) == \frac{\exp\left(\frac{F_h W_q \left(F_{a\&b} W_k\right)^T}{\sqrt{d}}\right)}{\sum \exp\left(\frac{F_h W_q \left(F_{a\&b} W_k\right)^T}{\sqrt{d}}\right)} \tag{14}$$

4.3 Cross-border trade export prediction based on improved reinforcement learning and BiGRU

After obtaining the multimodal data features affecting the sales volume of cross-border trade exports, this paper optimises the hyperparameters of the BiGRU model by using the MSRL proposed in Section 3, which efficiently searches for the optimal strategy automatically, improves the efficiency of hyperparameter searching, and then reduces the prediction error of BiGRU. The structure of MSRL-BiGRU is shown in Figure 5.

Firstly, the states $S = [s_1, s_2, s_3]$, s_1 indicates the total count of neurons in the BiGRU, s_2 represents the current count of BiGRU iterations, and s_3 represents the current studying rate of the BiGRU, and each state is defined as shown in equation (15).

$$\begin{cases} s_1 = \left[a_1^1, a_1^2, \ldots, a_1^n\right] \\ s_2 = \left[a_2^1, a_2^2, \ldots, a_2^n\right] \\ s_3 = \left[a_3^1, a_3^2, \ldots, a_3^n\right] \end{cases} \tag{15}$$

Each state action is denoted as $A = [s_1(A_1), s_2(A_2), s_3(A_3)]$, $A_i = [a_1, a_2, \ldots, a_n]$, and each state has n executable actions. The update of the Q value of the MSRL algorithm proposed in this paper is shown in equation (16).

$$Q(s, a) = Q(s, a) - \alpha\left[Q\left(s', a'\right) - Q(s, a)\right] \tag{16}$$

where $Q(s, a)$ is the Q-value of the current state and $Q(s', a')$ is the Q-value of the next state, the minimum Q-value is expected. The calculation of $Q(s', a')$ is as follows:

$$Q(s', a') = \begin{cases} R(s, a) - \gamma \max Q(s', a'), & \text{Not in the terminated state} \\ R(s, a), & \text{Reach termination state} \end{cases} \tag{17}$$

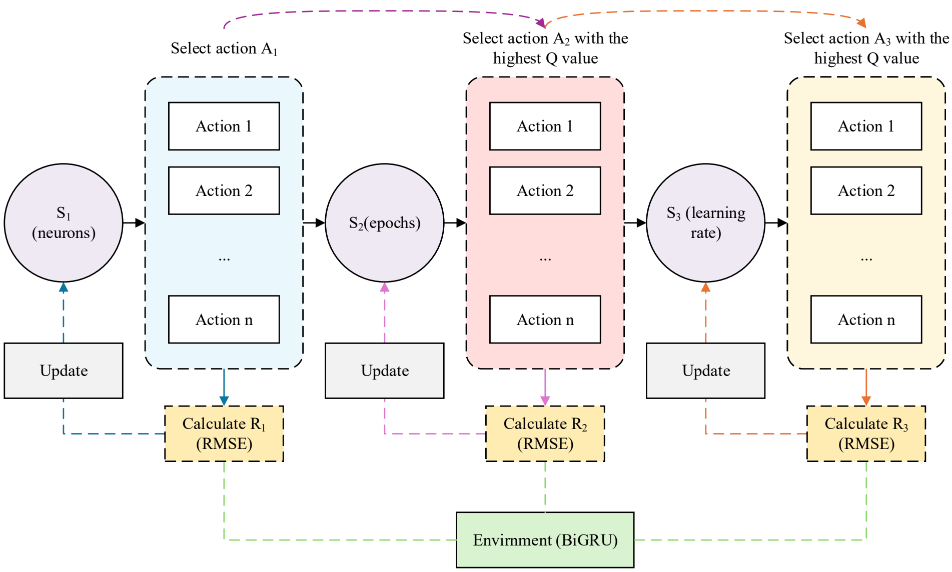
where $\gamma \in [0, 1]$ is the discount factor, $R(s, a)$ A signifies the reward value resulting from performing action a in state s to transfer to s', a' is the next action a selected by taking a greedy policy, and the reward operation is as below:

$$R(s, a) = RMSE(s, a) \tag{18}$$

where the value of $R(s, a)$ is the RMSE gained by performing a in s. $RMSE(s, a)$ is calculated as follows:

$$RMSE(s, a) = BiGRU(s, a) \tag{19}$$

Figure 5 The structure of MSRL-BiGRU (see online version for colours)



A BiGRU model is a GRU that consists of unidirectional, oppositely oriented GRUs whose network output is jointly determined by these two GRUs. In this section, a three-layer BiGRU network is constructed, and the expression of BiGRU is as follows, taking the L^{th} layer network calculation as an example.

$$h_t^L = f\left(W_{\vec{h}_t}^L \vec{h}_t^L + W_{\overleftarrow{h}_t}^L \overleftarrow{h}_t^L + b_t^L\right) \tag{20}$$

where h_t^L is the state of the obscured level at time t, b_t^L is the bias, and $\vec{h}_t^L$ and $\overleftarrow{h}_t^L$ represent the conditions of the forward and backward obscured levels, respectively, respectively, as follows:

$$\vec{h}_t^L = \text{GRU}^L\left(x_t^L, \vec{h}_{t-1}^L\right) \tag{21}$$

$$\overleftarrow{h}_t^L = \text{GRU}^L\left(x_t^L, \overleftarrow{h}_{t-1}^L\right) \tag{22}$$

where $\vec{h}_{t-1}^L$ and $\overleftarrow{h}_{t-1}^L$ are the forward and backward hidden layer states of the L layer at the time $t-1$, and x_t^L serves as the input for the L level.

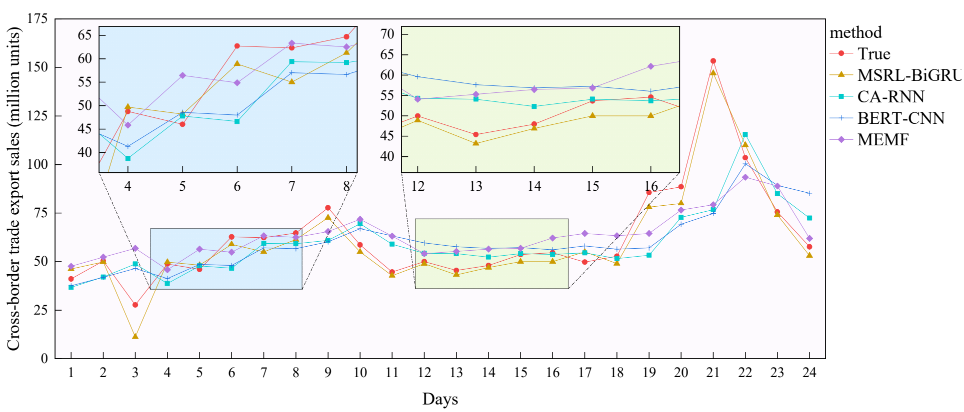
The input of MSRL-BiGRU is F^* and the output is the predicted value of export sales from cross-border trade for the BiGRU model. Initialise the Q-table by initialising the parameters of the MSRL, establishing each of the three optimisation parameters as three different states. Then, starting from state s_1, action a_1 is selected and executed, and the intelligent agent engages with the BiGRU's context to determine the RMSE value.

Monitor the reward function value alongside s', revise the Q table accordingly, and transition to the subsequent state s_2. When the final state s_3 is attained, the approach refreshes the state, and starts the cyclic search through s_1. The approach terminates the operation when the maximum amount of epochs is attained, so as to obtain the prediction value with the smallest RMSE, and to improve the prediction efficiency.

5 Experimental results and analyses

This section of experiments was conducted on a personal laptop using an 11th-generation Intel Core i7-11800H@2.30 GHz, 32 GB of RAM, an NVIDIA GeForce RTX 3060 GPU with 16 GB of video memory, and algorithms implemented in PyTorch. In model training, the discount factor was set to 0.1, the learning rate to 0.007, the batch size to 256, and the loss function to RMSE. The data used in this paper are from Alibaba AliExpress platform and Ministry of Foreign Trade and Economic Cooperation 'Internet Plus Foreign' from January 2015 to July 2016 Trade platform, China cross-border e-commerce integrated service platform, National Bureau of Statistics fixed product sales index statistics platform on leather shoes, machinery, electrical appliances and other manufacturing export products 7,294 sales detail data and 102,596 graphic comment data.

Figure 6 The results of predicting the export sales of cross-border trade (see online version for colours)



To verify the prediction performance of the proposed prediction method MSRL-BiGRU, three typical prediction methods BERT-CNN (An et al., 2023), CA-RNN (Cai et al., 2021), and MEMF (Xu et al., 2024) are selected for the comparison experiments in this paper. The results of the different methods for predicting the export sales of cross-border trade from 1 March to 24 March 2015 (denoted as 1, 2, …, 24) are shown in Figure 6. On 6 March, the real cross-border trade export sales were 612,000 units, and the predicted sales of BERT-CNN, CA-RNN, MEMF and MSRL-BiGRU were 469,000, 462,000, 547,000 and 589,000 units, respectively, and the predicted sales of MSRL-BiGRU were the closest to the real sales, while it can be seen from the overall trend that the MSRL-BiGRU's forecast results are more accurate.

In addition to analysing the results of the cross-border trade export sales forecasts of different methods, this paper also adopts MAE, MSE, RMSE, MAPE, and R^2, which are commonly used indicators to measure the forecasting performance, to assess the performance of each forecasting method, and the comparisons of MAE, MSE, RMSE, and MAPE of the four forecasting methods are shown in Table 1. The MAE and MAPE of MSRL-BiGRU were 0.1564 and 0.2306, which were reduced by 50.1% and 34.41% compared to BERT-CNN, 33.08% and 30.06% compared to CA-RNN, and 16.05% and 12.68% compared to MEMF, respectively. All the MSRL-BiGRU indicators were better than those of other comparative models. MSRL-BiGRU not only makes use of multilevel strategy to optimise the RL algorithm, but also extracts and fuses the features of static variables affecting export sales, text features of commodity reviews, and image features, which greatly enriches the multimodal feature representation and improves the prediction accuracy.

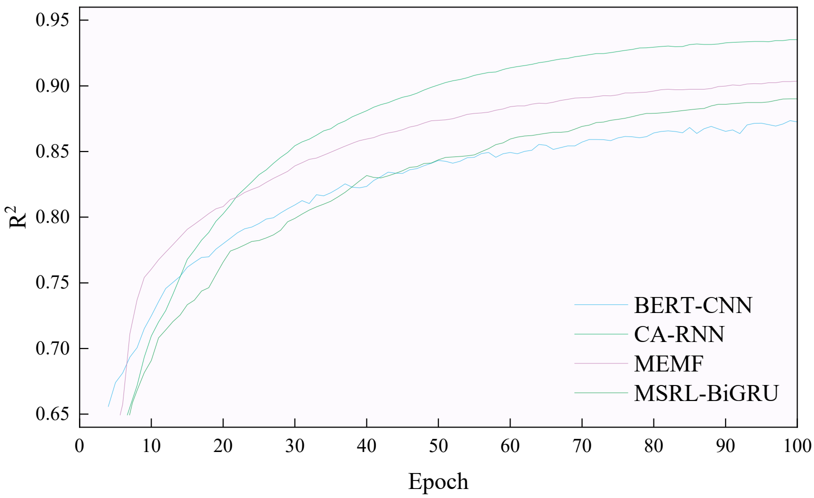
To visually verify the convergence performance of the different methods, training tests were conducted to compare the different models, and the optimal parameter values were uniformly adjusted to remain unchanged. The results of the comparison of the performance of the coefficient of determination R^2 of each method are shown in Figure 7. R^2 is an important index for assessing the prediction accuracy, and its value ranges from 0 to 1. The larger the value of R^2, the higher the prediction accuracy.

Table 1 Comparison of forecasting indicators for different forecasting methods

Method	*MAE*	*MSE*	*RMSE*	*MAPE*
BERT-CNN	0.3134	0.3896	0.5814	0.3516
CA-RNN	0.2337	0.3259	0.5085	0.3297
MEMF	0.1863	0.3061	0.4539	0.2641
MSRL-BiGRU	0.1564	0.2689	0.4168	0.2306

As the number of iterations increases, the convergence of the different methods increases and stabilises, with BERT-CNN performing relatively poorly and MSRL-BiGRU performing best. The converged R^2 of BERT-CNN, CA-RNN, MEMF and MSRL-BiGRU are 0.8028, 0.8716, 0.9043, 0.9527, respectively, and the R^2 of MSRL-BiGRU is the closest to 1, which indicates that it has the best prediction. BERT-CNN only considers a single textual modality, and the prediction results are worse than those of CA-RNN, MEMF and MSRL-BiGRU, which are multimodal prediction methods. Although CA-RNN considers image and text modalities, the RNN suffers from long-term dependency problems and overfitting, and does not optimise the prediction results, so the prediction performance is worse than that of MEMF and MSRL-BiGRU. MEMF also considers the multimodal influence characteristics of cross-border e-commerce export sales, but does not optimise the final prediction results, so the prediction accuracy is not as good as that of MSRL-BiGRU. Therefore, according to the experimental results, it is concluded that the learning ability of MSRL-BiGRU is stronger, and it can obtain a relatively high prediction accuracy within a limited number of iterations, and the amount of computation decreases with the number of iterations.

Figure 7 Determination R^2 of each forecasting method (see online version for colours)



6 Conclusions

With the acceleration of the process of global economic integration, cross-border trade occupies an increasingly important position in international trade. Accurate prediction of cross-border trade export sales is of key significance for enterprises to cope with market changes. This paper proposes a cross-border trade export forecasting method based on the fusion of RL and multimodal data to address the problem of insufficient single-modal feature extraction in existing studies, which leads to low forecasting accuracy. Firstly, the MSRL algorithm is designed, where the multilevel policy provides maximum confidence by jointly updating the semantic policy, and the multilevel reward jointly evaluates the visual-linguistic relevance by the visual reward to obtain the RL-specific target and improve the global decision optimisation capability. Then, CNN, Doc2Vec model, and improved ResNet152 model are used to extract static variable features, text features, and image features of cross-border trade export sales volume, respectively, and MMA is designed to meticulously capture and fuse the features of different modal data, and to distinguish between data of different semantic levels. Secondly, the fusion features are used as inputs to the MSRL-BiGRU model, and the three hyperparameters of the amount of neurons, epochs and studying rate of the BiGRU model are improved using MSRL, and each parameter to be optimised is established as a state, with a set of actions in each state, and the intelligences and the BiGRU continuously interact with each other in search of the optimal strategy, to obtain the optimal cross-border trade export sales prediction results. The proposed method significantly reduces the prediction error and shows excellent prediction results.

Although MSRL-BiGRU has been upgraded in the problem of cross-border trade export sales forecasting, the methodology of this paper is still deficient, and the subsequent research can be centred on the following points.

1 For the neural network hyper-parameter optimisation problem: after a hyper-parameter reaches the optimum, the MSRL algorithm updates the Q-function with the hyper-parameter that corresponds to the optimum at the next moment, and the parameters that are not sampled will not be selected as the optimum, so there is a situation that the optimum value will be missed. We will continue to refine the proposed methodology in future studies to obtain better learning strategies.

2 The work in this paper only considers three hyperparameters, namely, the amount of neurons, the amount of epochs of the model, and the studying rate, and subsequent work can expand more hyperparameters to find the optimal, such as: loss function, activation function, and the number of batch samples.

Declarations

This work is supported by the 2024 General Project of Zhejiang Provincial Department of Education (No. Y202454642).

All authors declare that they have no conflicts of interest.

References

An, W., Wang, L. and Zhang, D. (2023) 'Comprehensive commodity price forecasting framework using text mining methods', *Journal of Forecasting*, Vol. 42, No. 7, pp.1865–1888.

Cai, W., Song, Y. and Wei, Z. (2021) 'Multimodal data guided spatial feature fusion and grouping strategy for e-commerce commodity demand forecasting', *Mobile Information Systems*, Vol. 20, No. 1, pp.1–10.

Elia, S., Giuffrida, M., Mariani, M.M. et al. (2021) 'Resources and digital export: an RBV perspective on the role of digital technologies and capabilities in cross-border e-commerce', *Journal of Business Research*, Vol. 132, pp.158–169.

Gao, J., Li, P., Chen, Z. et al. (2020) 'A survey on deep learning for multimodal data fusion', *Neural Computation*, Vol. 32, No. 5, pp.829–864.

Jiang, J., Yao, W. and Li, X. (2021) 'Precise demand forecast analysis of new retail target products based on combination model', *Open Journal of Business and Management*, Vol. 9, No. 3, pp.1312–1324.

Kalaoglu, Ö.İ., Akyuz, E.S., Ecemiş, S. et al. (2015) 'Retail demand forecasting in clothing industry', *Textile and Apparel*, Vol. 25, No. 2, pp.172–178.

Kim, D., Seo, D., Cho, S. et al. (2019) 'Multi-co-training for document classification using various document representations: TF-IDF, LDA, and Doc2Vec', *Information Sciences*, Vol. 477, pp.15–29.

Lee, C-Y., Chou, B-J. and Huang, C-F. (2022) 'Data science and reinforcement learning for price forecasting and raw material procurement in petrochemical industry', *Advanced Engineering Informatics*, Vol. 51, p.101443.

Li, C. (2024) 'Commodity demand forecasting based on multimodal data and recurrent neural networks for e-commerce platforms', *Intelligent Systems with Applications*, Vol. 22, p.200364.

Lourenço, V.M., Pires, A.M. and Kirst, M. (2011) 'Robust linear regression methods in association studies', *Bioinformatics*, Vol. 27, No. 6, pp.815–821.

Massaro, A., Panarese, A., Giannone, D. et al. (2021) 'Augmented data and xgboost improvement for sales forecasting in the large-scale retail sector', *Applied Sciences*, Vol. 11, No. 17, p.7793.

Matsuo, Y., LeCun, Y., Sahani, M. et al. (2022) 'Deep learning, reinforcement learning, and world models', *Neural Networks*, Vol. 152, pp.267–275.

Menculini, L., Marini, A., Proietti, M. et al. (2021) 'Comparing prophet and deep learning to ARIMA in forecasting wholesale food prices', *Forecasting*, Vol. 3, No. 3, pp.644–662.

Ouyang, H., Wei, X. and Wu, Q. (2019) 'Agricultural commodity futures prices prediction via long-and short-term time series network', *Journal of Applied Economics,* Vol. 22, No. 1, pp.468–483.

Pan, H. and Zhou, H. (2020) 'Study on convolutional neural network and its application in data mining and sales forecasting for E-commerce', *Electronic Commerce Research*, Vol. 20, No. 2, pp.297–320.

Pham, H.T. and Yang, B-S. (2010) 'A hybrid of nonlinear autoregressive model with exogenous input and autoregressive moving average model for long-term machine state forecasting', *Expert Systems with Applications*, Vol. 37, No. 4, pp.3310–3317.

Saydam, S.G. and Civelek, M.E. (2022) 'Problems in cross-border e-commerce and development of cross-border e-commerce performance scale', *Journal of International Trade, Logistics and Law*, Vol. 8, No. 1, pp.133–140.

Wang, Y., Jia, F., Schoenherr, T. et al. (2020) 'Cross-border e-commerce firms as supply chain integrators: the management of three flows', *Industrial Marketing Management*, Vol. 89, pp.72–88.

Wang, Y., Wang, Y. and Lee, S.H. (2017) 'The effect of cross-border e-commerce on China's international trade: an empirical study based on transaction cost analysis', *Sustainability*, Vol. 9, No. 11, p.2028.

Xu, G., Ren, M., Wang, Z. et al. (2024) 'MEMF: multi-entity multimodal fusion framework for sales prediction in live streaming commerce', *Decision Support Systems*, Vol. 184, p.114277.

Zhang, L., Li, H., Zhu, R. et al. (2022) 'An infrared and visible image fusion algorithm based on ResNet-152', *Multimedia Tools and Applications*, Vol. 81, No. 7, pp.9277–9287.

Zhang, Y-D., Dong, Z., Wang, S-H. et al. (2020) 'Advances in multimodal data fusion in neuroimaging: overview, challenges, and novel orientation', *Information Fusion*, Vol. 64, pp.149–187.